

CSE 562 Midterm - SOLUTIONS

Question	Points Possible
A.1	5
A.2	10
A.3	20
A.4	15
B.1	10
B.2	10
C.1	10
C.2	10
C.3	10
Total	100

Part A. Relational Algebra, Optimization, and The ER Model
(50 points)

Assume we have two relations $R(A, B)$ and $S(B, C)$. All three attributes (A, B, C) are integer attributes. Assume that Relation R contains the following tuples: $(1, 2)$, $(2, 3)$, and $(3, 4)$. Assume that relation S contains the following tuples $(2, 2)$, $(2, 3)$, and $(5, 1)$.

1. (5 points) Give an example of an attribute (combination) that cannot be a key for relation S .

Answer:

B

2. (10 points) Which of the following relational algebra expressions are equivalent?

- | | |
|---|------------|
| (a) $\pi_A(\sigma_{B=1}(R))$ and $\sigma_{B=1}(\pi_A(R))$ | No |
| (b) $\pi_A(\sigma_{A=1}(R))$ and $\sigma_{A=1}(\pi_A(R))$ | Yes |
| (c) $\pi_{A, B}(R \times S)$ and $(\pi_{A, B}(R)) \times S$ | No |
| (d) $\sigma_{B=1}(R \times S)$ and $(\sigma_{B=1}(R)) \times S$ | No |
| (e) $\sigma_{B=1}(R \bowtie S)$ and $(\sigma_{B=1}(R)) \bowtie S$ | Yes |

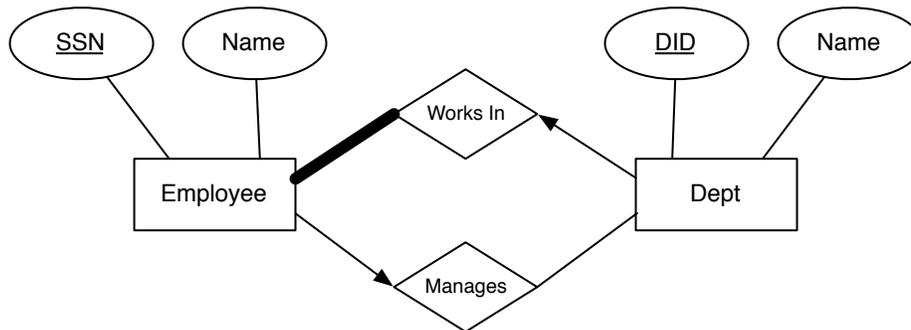
3. (20 points) Consider a relation $T(A, B, C)$. Express the relational division R/S using relational operators without the division operator.

Answer:

$$\pi_A(T) - \pi_A((\pi_A(T) \times S) - T)$$

4. (15 points) Draw an ER-diagram that captures the following two SQL statements

```
CREATE TABLE Employees(  
  ssn CHAR(11),  
  name CHAR(30),  
  did INTEGER,  
  PRIMARY KEY (ssn)  
  FOREIGN KEY(did) REFERENCES Dept,  
  ON DELETE CASCADE  
);  
CREATE TABLE Dept(  
  did  
  dname  
  mgrssn  
  PRIMARY KEY (did)  
  FOREIGN KEY (mgrssn) REFERENCES Employees(ssn),  
  ON DELETE NO ACTION  
);
```



Part B. Sql
(20 points)

Consider the following schema:

Suppliers(sid integer, sname string, address string)

Parts(pid integer, pname string, color string)

Catalog(sid integer, pid integer, price real)

1. (10 points) Describe in *one short* English sentence what the following SQL query computes:

```
SELECT P.pid, AVG(P.price)
FROM Parts P, Catalog C
WHERE P.pid = C.pid
GROUP BY P.pid
HAVING COUNT(*) > 1
```

Answer:

Find the PID and average price (1 point) for all parts (4 point) that have more than one supplier (5 point).

2. (10 points) Write a query *in SQL* to find the names of parts that are supplied by the most suppliers.

Answer:

```
SELECT P.PNAME
FROM PARTS P
WHERE P.PID IN (
  SELECT C.PID
  FROM CATALOG C
  GROUP BY C.PID
  HAVING COUNT(*) = (
    SELECT MAX(SUPPCUT)
    FROM (
      SELECT C2.PID, COUNT(*) AS SUPPCUT
      FROM CATALOG C2
      GROUP BY C2.PID
    ) PS
  )
))
```

Correctly selecting the PIDs that are supplied by the most suppliers can get 6 points.
Correctly selecting the PNAME through PID can get 4 points.

Part C. Indexes
(30 points)

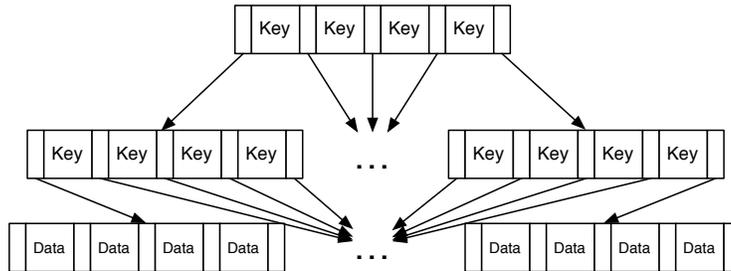
- (10 points) Describe a situation where you would build a B+Tree on the $\langle sid, pid \rangle$ attributes of the Catalog relation from Part B, and in no more than 3 sentences, contrast this situation with a situation where you would build a B+Tree on $\langle pid, sid \rangle$.

Answer:

$\langle sid, pid \rangle$ allows you to do range scans and enumerations of PID for specific SID values. Visa versa for $\langle pid, sid \rangle$.

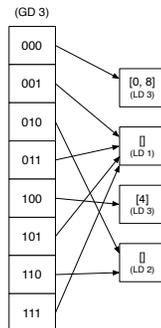
- (10 points) Consider that a B+Tree index has order $d=2$, and assume that the index has the following properties: (1) If I insert *any* additional data entry into the index, the height of the index would increase by one, and (2) The index has more than two index nodes. Draw a B+Tree index that has these two properties.

Answer:



- (10 points) Consider an initially empty extendible hashing index with a directory size of one, where the directory entry points to an empty bucket. Write down a sequence of three data insertions into this index such that the final index structure has a directory size of eight. You should also draw the final index structure after the insertions (including the directory, buckets, hash entries, local depths and global depth). You can assume that each bucket can hold at most two data entries.

Answer: Insert (using the identity as a hash function): 0, 4, 8



(Scratch Page)